PaVESy - User Guide

PaVESy - User Guide 1
Starting the application 3
Working with PaVESy 3
Creating database tables 6
Importing AraCyc flat file data 7
Importing KEGG Ligand data 8
Exporting database content 9
Importing database content 10
Creating folders and subsets 11
Pathway creation 14

Starting the application

The easiest way to work with PaVESy is to use Java[tm] Web Start Version 1.2. In this case you will work always with the latest software release of PaVESy. Details about the system requirements and installation instructions for Java[tm] Web Start can be found at Installation Guide Java [tm] Web Start. Starting our application via the Java[tm] Web Start link in our download page the software requirements will be tested automatically.
Another way to work with PaVESy is to download the jar-file. Windows and Mac OS X allow starting Java applications packed in jar-files by double clicking. If double clicking will not work follow the Linux instructions. The Linux instructions have to be accomplished inside a shell. Switch to the path where the PaVESy.jar file is stored and follow the instructions.
To start the application type:

user@linux: > **java -Xmx256m -jar PaVESy.jar**

If Java 1.4 or a higher version is installed and the JAVA_HOME and PATH variable is defined everything will work automatically. If an error occurs the JAVA_HOME must be defined to the path where it is installed, in our case **/usr/java/j2sdk1.4.0**.

Under Linux for example:

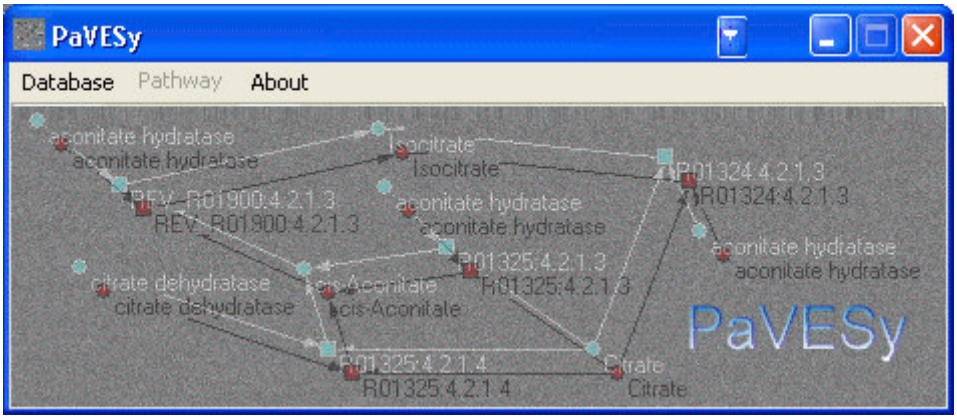
user@linux: > **export JAVA_HOME=/usr/java/j2sdk1.4.0**
user@linux: > **export PATH=$JAVA_HOME/bin:$PATH** *
user@linux: > **java -version**

*Windows user use a semicolon instead of the colon

If a declaration like java version 1.4... follows, everything is OK and you can start PaVESy again.

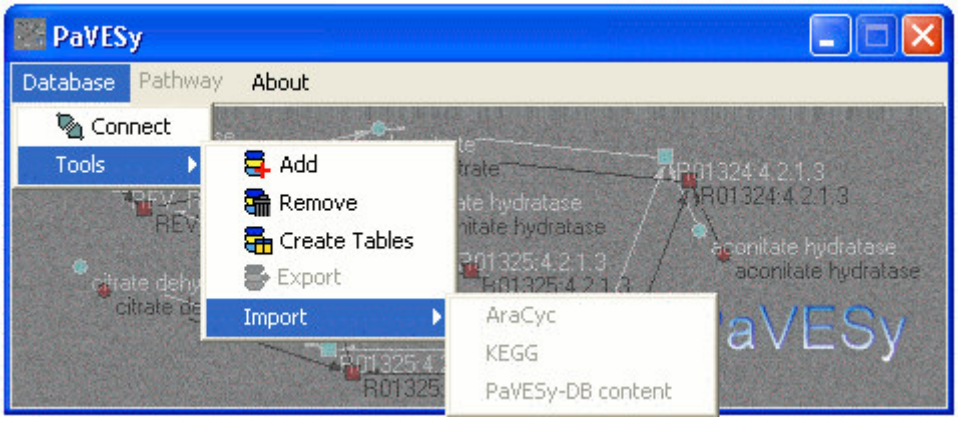
Working with PaVESy

After starting the application the main window (**Picture 1**) offers the possible functions of PaVESy.

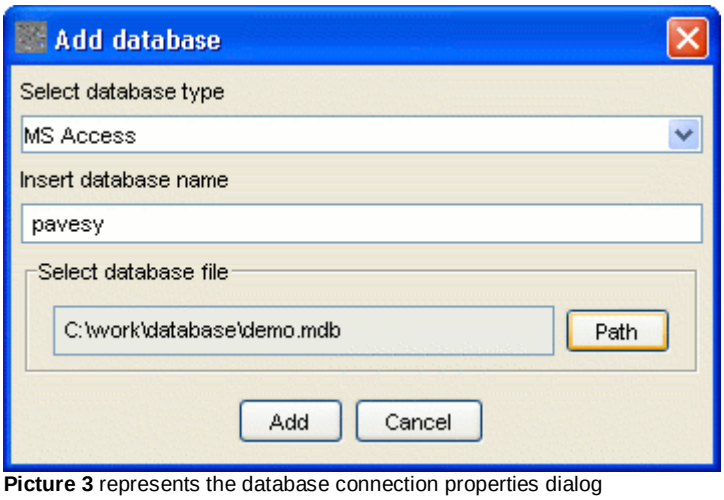


Picture 1 shows the main window of the application

In this window you have the possibilities to choose between two menus. The database menu which contains a submenu called Tools to define database connections, delete these connections, creating tables in empty database instances and if you are connected to a database the import and export functions for complete databases. The second option is the *About* menu to get information about PaVESy. If a database connection is defined and you selected this connection as working connection the Pathway menu is activated and you can use the database explorer, pathway builder and search or visualization tool. Before connecting to a database is allowed a database destination must be defined via the *Add*-Dialog (**Picture 3**) of the database submenu Tools displayed in **Picture 2**.

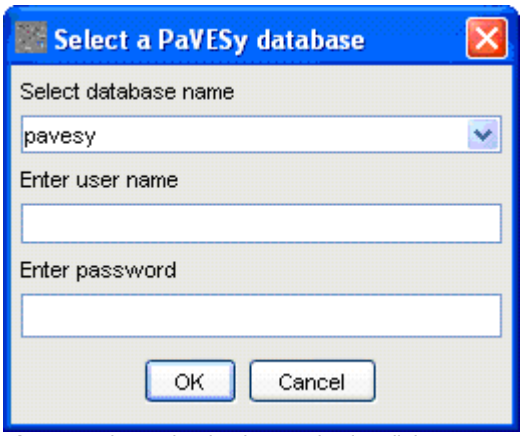


Picture 2 displays the menu to define database settings



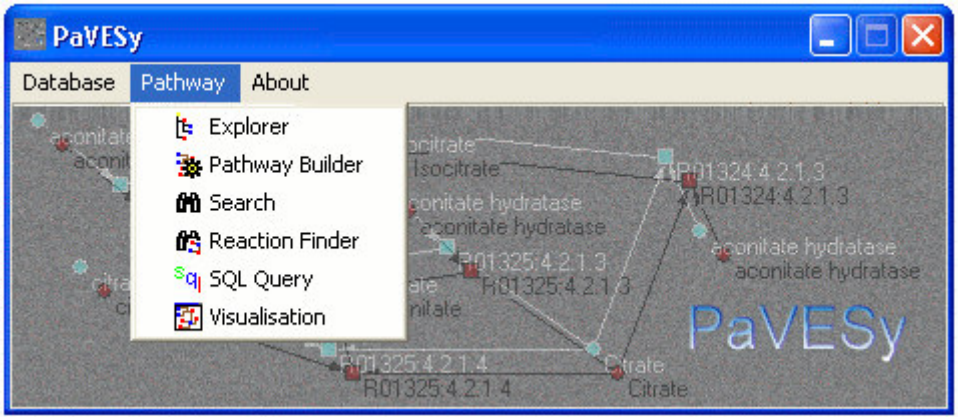
Picture 3 represents the database connection properties dialog

Picture 3 shows the settings for a MS Access data source. The user has to define a name and the path to the MS Access file. Adding a SAP DB needs the URL where the database is located. If these settings are confirmed, it is possible to choose this data source via the *Connect*-dialog (**Picture 4**). Database resources for connections can be deleted by calling the *Remove*-dialog over the *Remove* entry in the Database menu.



Picture 4 shows the database selecting dialog

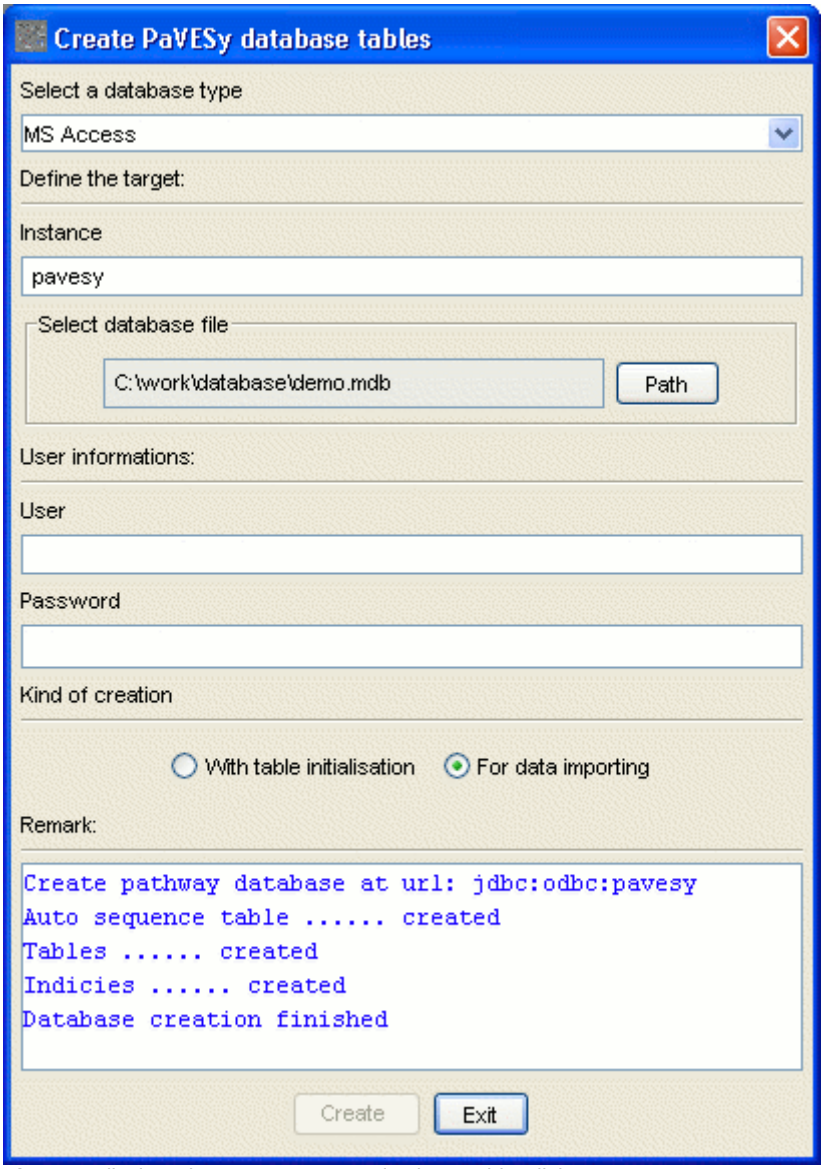
After validation of selection the *Pathway* menu seen in **Picture 5** is enabled and the mentioned functions are available.



Picture 5 pictures the menu to access the PaVESy main functions

Creating database tables

Storing and editing data for PaVESy needs a database connection to a defined database source. In this source, the data must be stored in a defined table organization. To simplify the table creation the *Create Tables* menu item is include in the application. Choosing this menu item calls the *Create PaVESy database tables* dialog displayed in **Picture 6**.

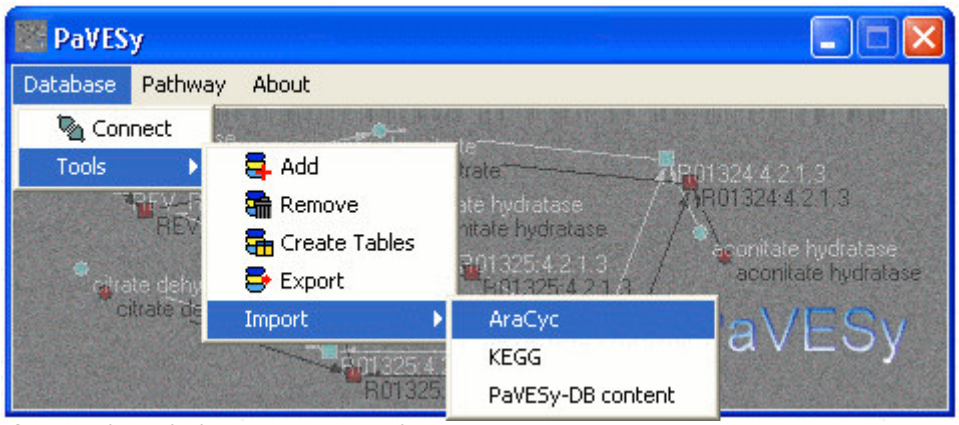


Picture 6 displays the *Create PaVESy database tables* dialog

For table creation different information is needed. At the moment two database types are supported, the SAP DB or MS Access. Using MS Access, PaVESy has to run under Microsoft Windows, because PaVESy uses the standard ODBC driver. Further particular is the name of the instance. Using MS Access the path to the data file is also essential. In the case of using SAP DB the URL to the database instance must be defined. At last the user can choose between an empty or initialized database. For data importing the database has to be empty. Initialized means that the database contains object classes like compounds, reactions, pathways, subsets and so on.

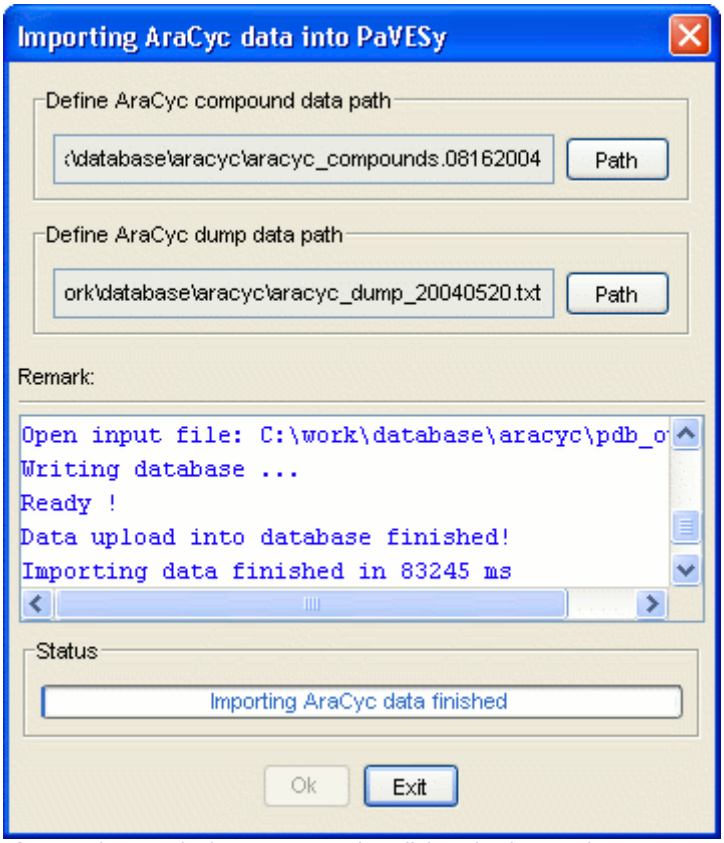
Importing AraCyc flat file data

The AraCyc importing tool is available via the *Tools* submenu *Import* (**Picture 7**) located in the database menu. Before using this function a database destination must be selected by using the *Connect* option.



Picture 7 shows the import AraCyc menu item

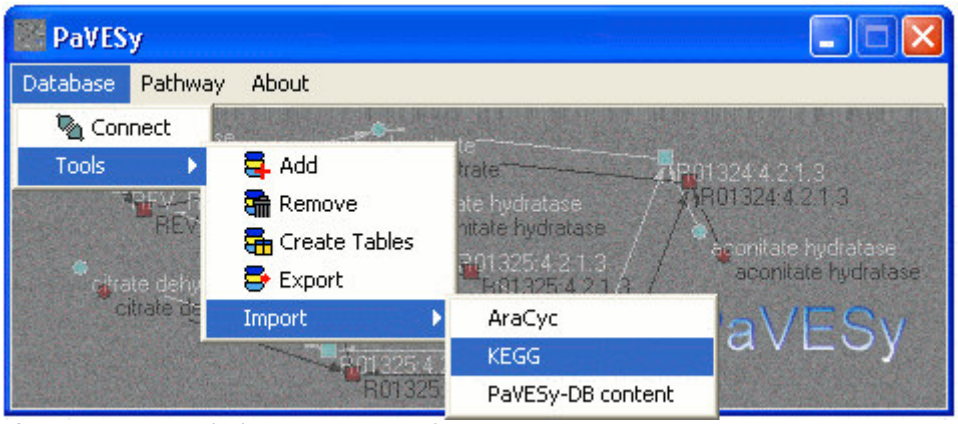
Before importing AraCyc data it is necessary to define the path to the AraCyc compound and dump file. The *aracyc_compounds* and *aracyc_dump* file can be downloaded from the TAIR ftp server. Defining the paths to the files can be done by using the import dialog, displayed in **Picture 8**. After defining the paths, the import function can be started by clicking the *Ok* button. The white panel contains all steps of importing which are finished.



Picture 8 images the import AraCyc data dialog after importation

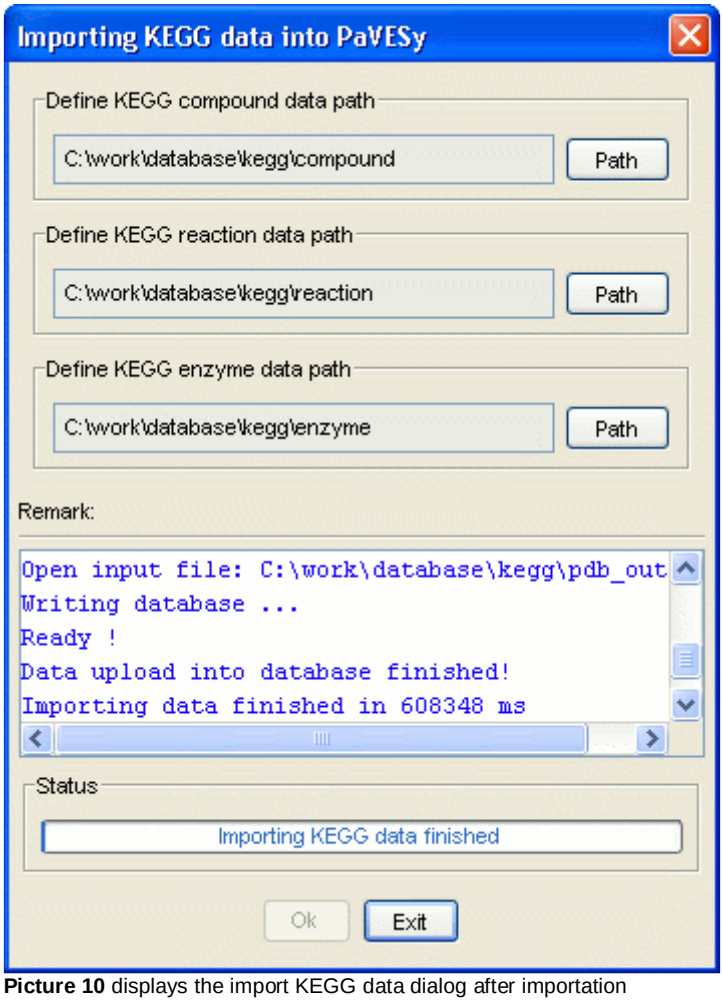
Importing KEGG Ligand data

The KEGG importing tool is available via the Tools submenu *Import* (**Picture 9**) located in the database menu. Before using this function a database destination must be selected by using the *Connect* option.



Picture 9 represents the import KEGG menu item

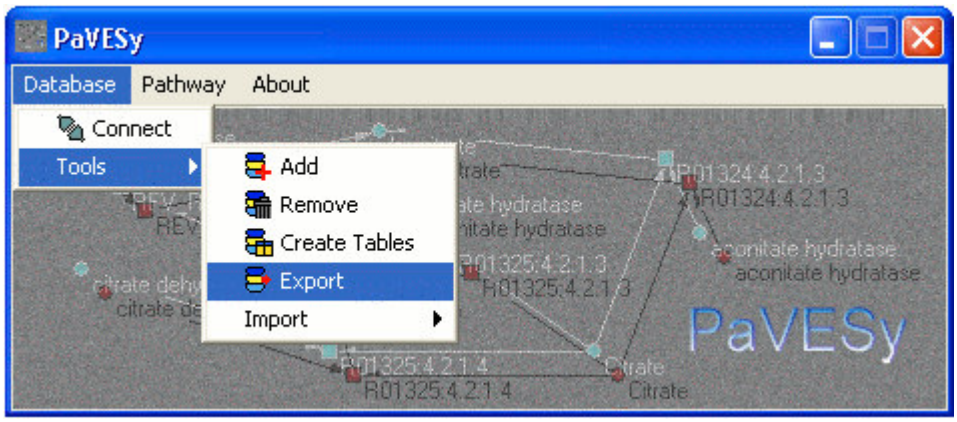
Before importing KEGG data it is necessary to define the path to the KEGG Ligand compound, reaction and enzyme file. The files are available on the KEGG ftp server and are packed into the *ligand.weekly.last.tar.Z* file. Defining the paths to the files can be done by using the import dialog, displayed in **Picture 10**. After defining the paths, the import function can be started by clicking the *Ok* button. The white panel contains all steps of importing which are finished.



Picture 10 displays the import KEGG data dialog after importation

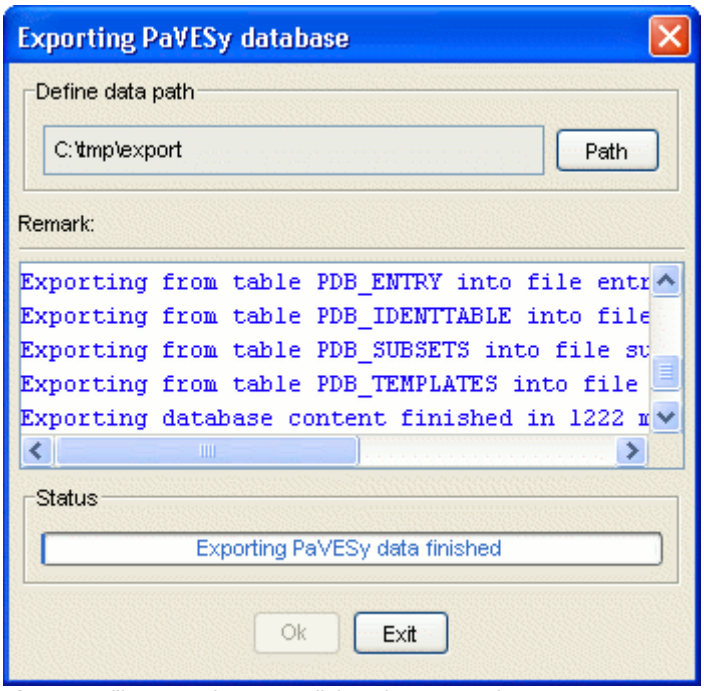
Exporting database content

The exporting tool is available via the database submenu *Tools* (**Picture 11**). Before using this function a database destination must be selected by using the *Connect* option.



Picture 11 shows the export menu item

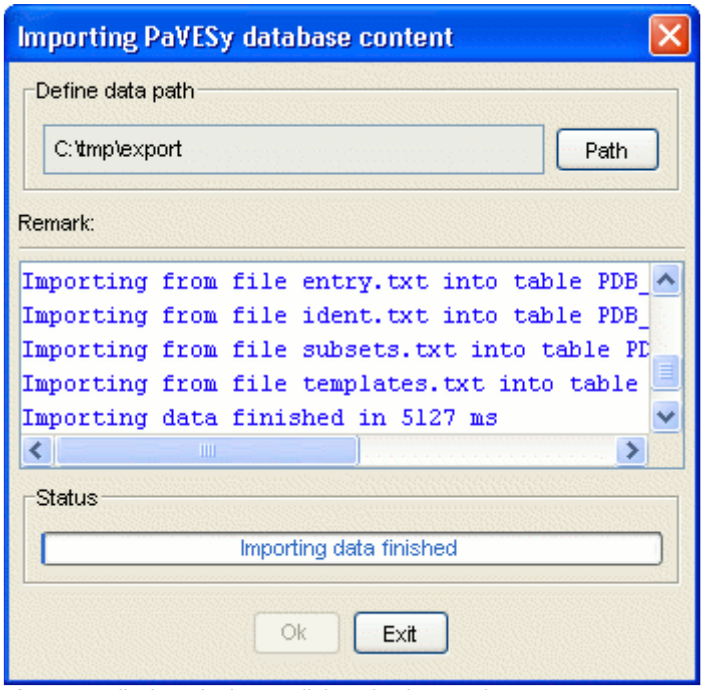
In the export dialog, displayed in **Picture 12**, the user has to choose the path where the gzip text files which contain the table data should be stored. Defining the path can be done by pressing the *Path* button in the dialog. After defining the path, the export function can be started by clicking the *Ok* button. The white panel contains all steps of exporting which are finished.



Picture 12 illustrates the export dialog after exportation

Importing database content

Importing data needs the same connection conditions as exporting. Furthermore files are needed which contains the data to import. To start the import tool, selecting the *Import* submenu item seen in **Picture 11** is necessary. In this submenu three import options are available. First, data import from AraCyc flat files. Second, importing data from KEGG Ligand files and last but not least the PaVESy database content import. For importing PaVESy database data, it is possible to use the gzip text files from our download side. After selecting *PaVESy-DB content* option the import dialog presented in **Picture 13** opens.

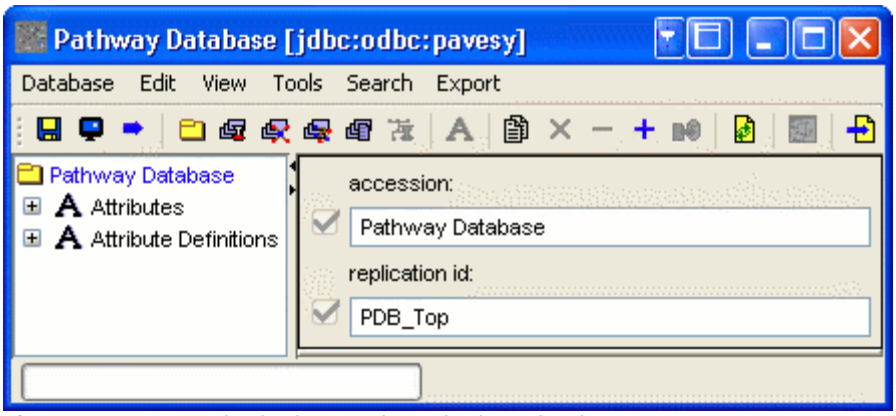


Picture 13 displays the import dialog after importation

The import dialog consists of two parts. First part where to define the path of the files which contains the data to import. Second the display part, where the import status is indicated. After defining the data source path, the *Ok* button is activated and the import can be started. A precondition for importing data is an empty database instance with all needed tables. Defining these tables can be done by the *Create tables* option in the *Tools* submenu.

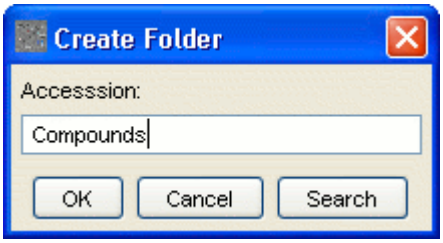
Creating folders and subsets

After importing data the database explorer shows no database entries like compounds, reactions or pathways.



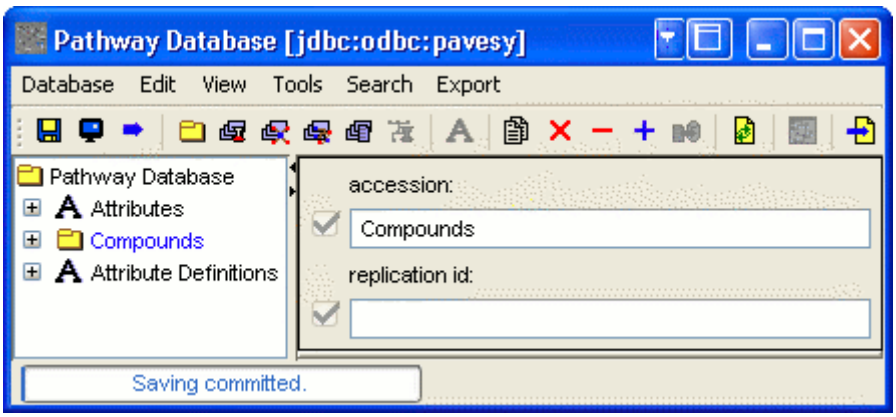
Picture 14 represents the database explorer after importing data

To display these entries it is necessary to create folders and subsets which contain the database elements. These folders allow structuring the database content in a user driven way.
A folder could contain subfolders and subsets. A subset holds compounds, reactions or pathways. A mixture of compounds, reactions or pathways in a subset is not possible.
For creating a folder pressing the yellow folder button in the button bar is essential. The last step before creating the folder is to define the name in the dialog.



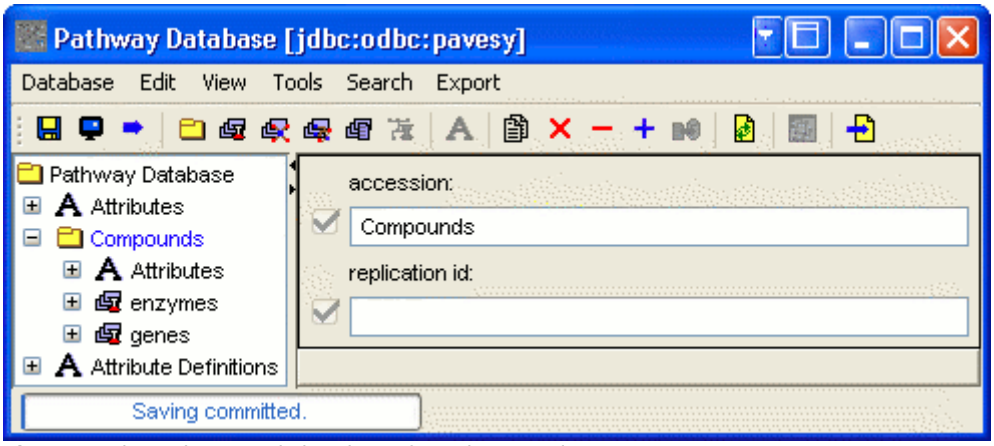
Picture 15 displays the folder creation dialog

To create subfolders selecting the parent folder is inevitable, because the new folder will be added into the selected layer. If no layer is selected the folder will be created in the top level layer. After folder creation a refresh of the explorer is essential to show the new folders. This can be done by the *View* menu or the *Refresh* button.



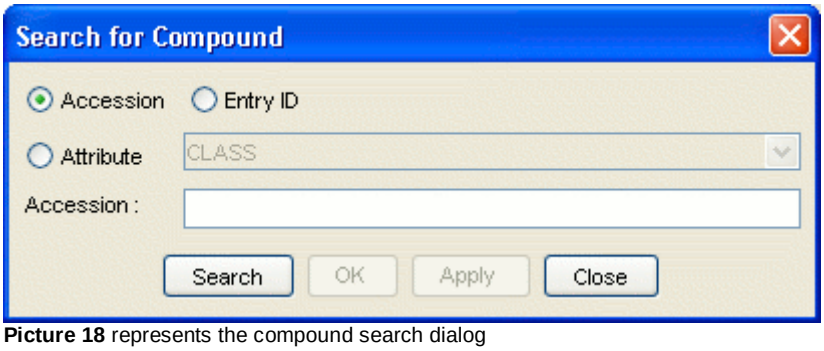
Picture 16 displays the refreshed explorer after folder creation

Subsets allow displaying the elements of the database in the explorer. Different types of subsets are available: Compound-, reaction- and pathway-subsets.
Compound- and reaction-subsets simplify the pathway creation. The creation will be explained later. In our example, we create compound-subsets, to separate enzymes and genes into different subsets.
At first we create an enzymes- and genes-compound-subset by using the *Create new compound subset* button next to the *Create new folder* button. The subsets should be stored in the *Compound* folder seen in **Picture 16**.
After creation and opening the *Compound* folder by expanding the tree, the new subsets will be imaged.



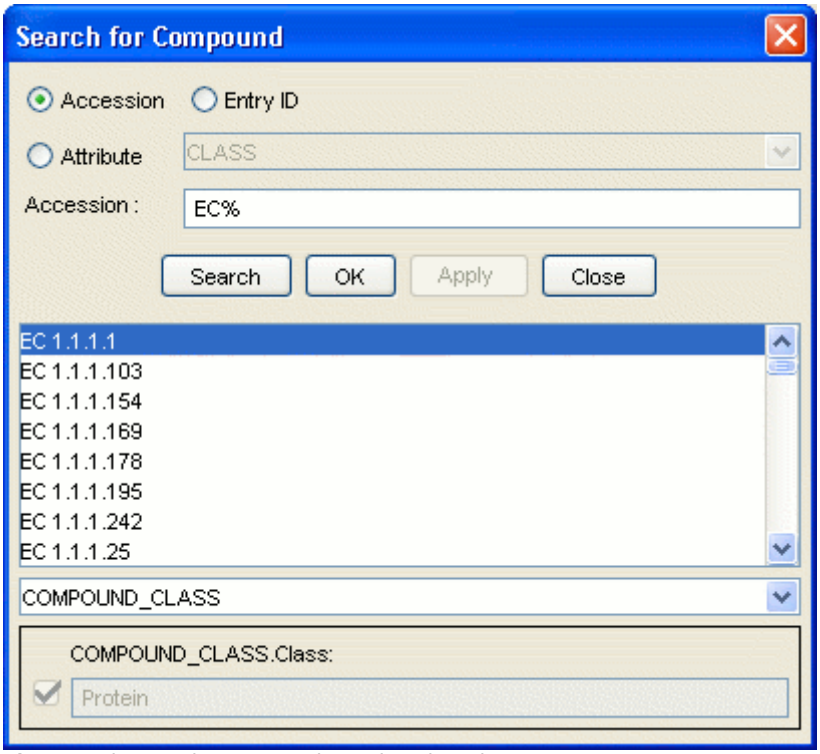
Picture 17 shows the expanded explorer after subset creation

Adding all enzymes of the database to the *enzymes* subset a database search is required. The blue *plus* in the button bar allows adding of database entries. This button opens a type depending database search dialog shown in **Picture 18**.
In the case of the *enzymes* subset, only compounds are allowed to add.



Picture 18 represents the compound search dialog

The *Search for Compound* dialog allows to search for defined entry accessions, entry ids or defined attributes. In the case of adding enzymes to the subset, we search for an accession which starts with *"EC"*.
To define wild cards for the following letters we use the percent sign. The results of the search will be displayed in the lower pop up part of the dialog.

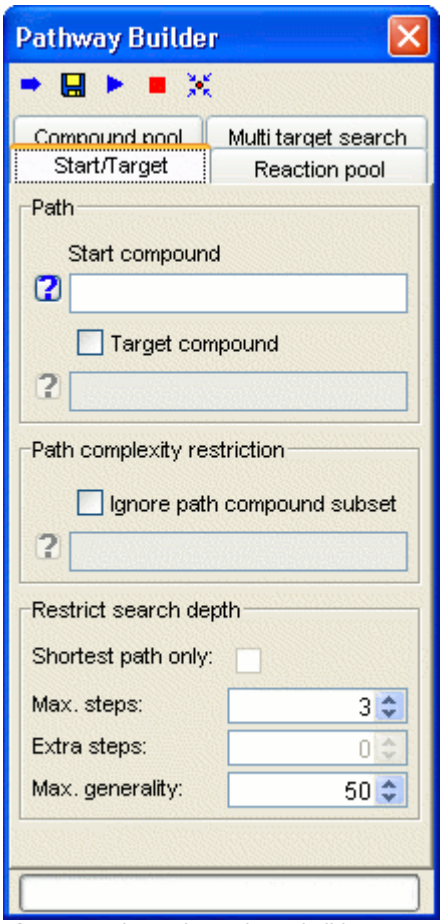


Picture 19 images the compound search and result

For every result it is possible to show defined attributes by selecting the attribute name via the combo box.
For adding the enzymes to the subset select them in the list and press the *OK* button and approve the following question with *Yes*.
The same can be done for the genes by entering *"At%"* in the *Accession* field of the *Search for Compound* dialog seen in **Picture 18**.

Pathway creation

Before starting the pathway creation a subset for the pathways must be generated, because the Pathway builder needs a subset for storing the constructed pathway. Subset and folder creation can be found in the Creating folders and subsets part.
Creating pathways is possible via the Pathway builder displayed in **Picture 20** from the *Pathway* menu of the main window (**Picture 5**). Starting the Pathway builder delivers different options for creation. The options will be explained in the following part.



Picture 20 shows the Pathway builder

The Start / Target search

The user is able to define a start compound by clicking the question mark left of the *Start compound* text field displayed in **Picture 20**. It is the same procedure as defining compounds for a subset. Defining a target compound and ignore compound subset is optional. Declaring a target the builder search for path between the source and target compound but is restricted by the maximum number of steps. This number defines the number of reactions which are placed between these two compounds. Choosing the *Ignore path compound subset* option defines compounds which should be ignored by pathway creation. This would be interesting for compounds like ADP or H2O or compounds which not interesting for later work. These restrictions are available for most of the building options.

Multi target search

It works like the start target search but more then one target has to be defined. The target compounds must be stored in a subset which is predefined in the database.

Compound pool

In this option, selecting a compound subset is necessary. The builder tries to complete the paths between the compounds of the subset.

Reaction pool

The reaction pool option is abutted on the compound pool option, but in this case a reaction subset has to be selected and the path between these reactions where completed.

After starting the creation by pressing the *blue arrow* the builder starts. Increasing the maximum number of steps means rising up the time of creation. If the creation is finished, a window opens where the user has to select the pathway subset in which the pathway should be stored. After defining a name for the pathway pressing the *store* button is necessary. Subsequently closing the *Pathway builder* and switching to the *Database explorer* is possible. Selecting the created pathway in the tree of the explorer, on the right side the elements of the pathway where listed. The pathway can be visualized in the integrated *Visualizer* by clicking *Open the Visualizer* button. In this window different layout functions are included. Moving the edges, get detailed information or adding and deleting of reactions is possible.

For critics, wishes or proposals please contact us.